NASA DEVELOP National Program
Pop-Up Project – Hunter College

Summer 2024

Central Park Ecological Conservation
Assessing Tree Health Conditions in New York City's Central Park with NASA Earth Observation Data

DEVELOP Technical Report

August 9th, 2024

John Hocknell (Project Lead)
Jack Graziano
Ikram Morso
Kayla Paramore

Advisors:
Dr. Mehdi Heris, City University of New York Hunter College (Science Advisor)
Joseph Spruce, Analytical Mechanics Associates (Science Advisor)

Lead:
Ella Haugen (Virginia – Langley)

1. Abstract

The Central Park Conservancy stewards New York City's iconic Central Park with a mission to preserve the park for all. This mission is complicated by the spread of Dutch elm disease (DED) which has threatened the culturally and ecologically significant American elm tree (*Ulmus americana*). Central Park is home to one of the largest and last remaining urban concentrations of American elm and the Conservancy currently protects them through integrated pest management. This paper discusses an interdisciplinary feasibility study that assessed the application of NASA Earth observations from 2014 to 2023 to detect changes in forest phenology possibly related to DED. Landsat 8 and 9 imagery was used to calculate multiyear time series of the Normalized Difference Vegetation Index (NDVI) and quantify changes in land surface phenology for a given year. A pixel-based logistic regression analysis was performed using changes in NDVI, tree site locations, and recorded occurrences of trees infected with DED as inputs. The results of this analysis show that changes in NDVI derived from Landsat data are capable of detecting unhealthy tree canopies with 71% precision and healthy tree canopies with 41% precision. The study had uncertainties and limitations due to the spatial and temporal resolutions of Landsat, the natural variability in land surface phenology and NDVI, and the attempt to detect disease impacts while disease prevention and mitigation is occurring. As is, the findings of this study and its methods provide managers with an approach for integrating Earth observations to make more informed decisions in the application and timing of urban forest management activities.

Key Terms

Central Park, remote sensing, urban forest management, pest management, Landsat, Normalized Difference Vegetation Index (NDVI), American elm, Dutch elm disease, vegetation phenology

2. Introduction

2.1. Background Information

With more than 160 years of history, Central Park is an iconic part of New York's cityscape (Central Park Conservancy, 2024). Its urban forest of more than 18,000 trees is composed of 170 species and is a vital component of the city's green infrastructure (Broderick, 2024). Healthy forest ecosystems can cool peak summer temperatures, absorb and filter stormwater, reduce air pollution, release oxygen, store carbon in vegetation and soils, and support biodiversity (Bounds et al., 2014). In contrast, degraded and unhealthy forests have a diminished capacity to deliver these services (Li et al., 2019). There are many challenges in maintaining a vibrant ecosystem in a dense, city environment. Abiotic stressors like climate change and increasing urban heat, severe storms, and constant heavy foot traffic all add to the complexity of maintaining a healthy urban forest within the park. Biotic stressors such as invasive species, insect pests, and diseases also can place enormous pressure on tree health (Bounds et al., 2014; Staley, 2022; Broderick, 2024). The American elm tree (*Ulmus americana*) serves as a prime example of a species threatened by these three biotic stressors and their interactions. This culturally significant tree species once lined the streets of many American cities, made popular by the shade provided by their grand canopies, but is now greatly reduced in occurrence due to the spread of an invasive fungus (*Ophiostoma ulmi*) that causes Dutch Elm Disease (DED). This fungus spreads by insect vectors such as the elm bark beetle *(Scolytus multistriatus)* and directly between trees through root grafts (Copeland, 2022). DED spreads and kills elms quickly. It can be a matter of days from the first signs of flagging to the point when the tree is no longer salvageable, making DED monitoring and prevention especially difficult and crucial (Broderick, 2024).

2.2. Partner Concerns and Objectives

Mitigating the impacts of American elm stressors is an essential part of the Central Park Conservancy's mission. For more than 40 years, the Conservancy has been the official caretaker of the park to ensure that it continues to provide crucial community and ecosystem services (Central Park Conservancy, 2024). Central Park has one of the largest and last remaining urban forest collections of American elm trees in the world. The Conservancy is particularly concerned about the approximately 1,600 American elms in the park due to the prevalence of DED (Broderick, 2024). Within the 843-acre park are four zones of interest due to their high concentrations of American elm trees: The North Woods, The Ramble, The Mall, and the west side of 5th Avenue (Figure 1).

The Mall and the west side of 5th Avenue are areas of high pedestrian traffic, whereas The North Woods and The Ramble are more densely forested and considered natural areas that together make up 76 acres of woodlands within the park (Sutliff, 2023).

Figure 1. Map of Central Park (outlined in black) with areas of relatively high concentrations of American elm trees (outlined in pink).

The Conservancy currently protects American elms through integrated pest management and site inspection for diseased trees. Their forest management plan follows a comprehensive cyclical approach that operates over 7-year periods. Certified arborists assess each tree in the park through regular inspections and biological sample testing to determine the risks to a tree and how often it must be re-inspected. Special projects in the park, including integrated pest management tasks, require increased coordination and precise timing to ensure their efficacy. The Conservancy takes a proactive approach to this process by predicting when these special projects will occur based on previous projects and identifying trends in the timing of cyclical biological events (phenology) of trees and pests attacking trees.

The Conservancy partnered with NASA DEVELOP to learn how satellite remote sensing can be integrated into its forest management planning. With this integration, the Conservancy aspires to make more informed decisions about the planning and timing of its tree inspections and special projects. Our team's primary objective was to assess the feasibility of using NASA Earth observations of Central Park to quantify changes in the health of American elm trees from 2014 to 2023. More specifically, we aimed to evaluate the detection of DED at the pixel level using Landsat 8 and 9 imagery and recorded occurrences of trees infected with the disease within the park. We also sought to analyze how temperature and precipitation play a role in assessments of tree foliage phenology using data from a weather station located in the park.

2.3. Scientific Basis

Vegetation phenology observed from remote sensing, also known as land surface phenology (LSP), is commonly estimated using vegetation indices derived from the red and near-infrared regions of the electromagnetic spectrum as a proxy (Caparros-Santiago, 2021). The Normalized Difference Vegetation Index (NDVI) is derived from these regions of the spectrum and is the most commonly used index when mapping vegetation (Neyns & Canters, 2022). Analyzing seasonal changes in vegetation phenology and NDVI can help identify and assess forest threats due to insect damage (Hargrove et al., 2009). Understanding the health of trees and the impact of diseases necessitates studying pest development and shifts in Growing Degree Days (GDD) for proactive management (Dean & Hodgson, 2022). GDD values, which measure heat accumulation critical to insect life stages, are essential tools in Integrated Pest Management (Murray, 2020). Following this basis, we utilized a diverse range of data such as NDVI, GDD, and weather data.

3. Methodology

3.1 Data Acquisition

3.1.1. Satellite and LiDAR data

We acquired multispectral satellite imagery from Landsat 8 and 9 for our entire study period (Table 1). We utilized the classified land cover derived from the LiDAR point cloud to extract a multi-polygon tree canopy layer used to monitor the tree canopy using remote sensing, excluding areas without trees. These excluded areas include water bodies, buildings, and construction sites.

3.1.2. In situ Data

The Central Park Conservancy provided us with three datasets. The first dataset held information about tree location, tree species, and the history of the tree. The second dataset contained information regarding tree location, dimensions (diameter at breast height, crown width), and management activities related to individual trees. The final set of data provided by the Conservancy contained zonal shapefiles outlining specific areas within the park. We acquired weather station data from 2000 to 2023 using the Global Historical Climatology Network-Daily (GHCN) database. This database included daily temperature maximums and minimums, as well as daily precipitation.

Table 1
Remotely sensed data

Satellite	**Sensor**	**Product**	**Spatial Resolution**	**Temporal Resolution**	**Reference**	**Acquisition method**
Landsat 8	Operational Land Imager (OLI)	Spectral Radiance (13 bands)	30 meters	16 Days (Solo) 8 Days (for years where Landsat 8 and 9 data are present)	(Landsat 8, 2021)	Google Earth Engine Catalog

Landsat 9	Operational Land Imager 2 (OLI-2)	Spectral Radiance (13 bands)	30 meters	16 Days 8 Days (for years where Landsat 8 and 9 data are present)	(Landsat 9, 2021)	Google Earth Engine Catalog
LiDAR	Light Detection and Ranging (LiDAR)	Tree canopy derived from LiDAR Point Cloud (10 classes)	1-foot derived Bare Earth DEM	1 product (2017)	(Office of Technology and Innovation, 2022)	Discover GIS Data NY

3.2 Data Processing

3.2.1. Satellite and LiDAR data

Firstly, we used land cover data derived from LiDAR to extract Central Park's tree canopy. This was then intersected with shapefiles provided by the Central Park Conservancy to create zonal tree canopy polygons. We processed Landsat 8 and 9 imagery using JavaScript within the Google Earth Engine API. We filtered through the Landsat 8 and 9 collections selecting all dates with less than 30% cloud cover between January 2014 and December 2023. We also utilized Landsat's QA layer to mask out areas occluded by clouds. We used JavaScript to compute a Normalized Difference Vegetation Index (NDVI) band that was stacked with the Landsat spectral bands for a given acquisition date.

3.2.1. In situ Data

In preparation for the validation step, we created an elm tree mask to select pixels with elm tree occurrences and exclude all other pixels. We built the mask by creating radius buffers around elm tree points provided by the Conservancy, based on each tree's canopy diameter. Each tree point record had three fields: tree type, tree diameter at breast height (DBH), and tree canopy diameter, but some of the records were missing information on DBH and canopy width. We filled in the missing canopy data using a two-step process. First, we ran a linear regression on the approximately 54.14% of elm points that had both DBH and canopy data. We found that the R^2 value of 0.61 indicated a useful correlation for estimating missing data, so we applied the equation y = mx +b where x was the DBH, m was equal to 1.2739, and b was equal to 11.8972. This method filled in the estimated 31.64% of the elms that had DBH, but no canopy spread. For the 14.22% of the records that had neither DBH nor canopy spread data, we took the average of the canopy spread data, which was 30 feet and applied it to the missing records.

3.2.2 *In situ Data for Validation*

Furthermore, we utilized in situ data to validate our pixel-level NDVI data. This involved joining three datasets: tree activities, tree sites, and pixel-level NDVI. These datasets were joined according to their geospatial coordinates. The coordinates for all in-situ data utilized the WGS84 coordinate system and were rounded to the fourth decimal place since the coordinate systems varied in precision. After joining the datasets, pixels that contained a tree with DED were marked with a "True" value, and pixels that did not contain DED-infected trees were assigned a "False" value.

3.2.3 *In situ Weather Data*

In addition to processing Conservancy data, the team also processed precipitation and temperature data. We first calculated the daily mean temperature to determine the growing degree days GDD (equation 1), where GDD_i is the cumulative growing degree days accumulated from March 1st to day *i*, with *i* representing the number of days after March 1st, and T_{max} and T_{min} represent the daily recorded minimum and maximum temperatures, respectively (Equation 3). Elm bark beetles are most active at approximately 700 cumulative degree-days above 52°F, accumulated from March 1, which clarifies that the base temperature (T_{base}) in the equation is 52°F, below which elm bark beetles do not feed, grow, or reproduce (Dreistadt & Lawson, 2014).

$$GDD_i = \Sigma_1^i \frac{(T_{max,i} - T_{min,i})}{2} - T_{base} \tag{1}$$

Secondly, we calculated the average precipitation, cumulative growing degree days (GDD), and cumulative precipitation from January until the full spread of beetles, which we identified by the first tree removal dates each year from 2015 to 2023. This analysis allowed us to track the relationship between these variables and beetle spread.

3.3. Data Analysis

3.3.1. *Land Surface Phenology*

We used Landsat 8 and 9 imageries to compute the NDVI for each date. NDVI is widely used to detect the start of the vegetation season due to its ratio-based properties that can reduce influences from sun angles, topography, clouds, and atmospheric conditions. We calculated yearly mean NDVI values for each index to better characterize the overall American elm population and enhance anomaly detection. Subsequently, we created graphs illustrating changes in tree canopy from 2013 to 2023. Using the temporal analysis using the NDVI values, we determined the start and end of the season. These NDVI graphs of the tree canopy allowed us to identify approximate season start and end dates, years with high mean NDVI, months showing peak vegetation, and insights into tree growth and greenness throughout the study period. (Banskota et al., 2014).

We accomplished this step using Generalized Additive Model (GAM) analysis, which calculates the fitted NDVI values using the approximate derivative method (Guo et al., 2021). This method defines a threshold for significant change (the median of the absolute derivative) to determine the start and end dates of the season. From these dates, we extracted the days of the year corresponding to the start and end of the season from the GAM analysis output. This result helps in understanding the changes in seasonality each year in the study area. To identify the vegetation peak during the growing season each year, we calculated the maximum NDVI and extracted the corresponding dates for each year.

3.3.2. Tree Canopy Change

As an exploratory tool for observing changes in the tree canopy NDVI, we produced a map showing deviations from the 10-year tree canopy average for each year of the study period. To make these maps, we retrieved the Landsat 8 and 9 images with less than 30% cloud cover within the vegetative season (April 1 – Sept 30) for all years available (Jan 1, 2013 to July 30, 2024, for a period of ten years plus two partial vegetative seasons in 2013 and 2024). We took the mean values of these images to produce an average 10-Year Vegetative Growing Season NDVI Map. Next, we created a mean NDVI map for each year's vegetative season, respectively. Finally, we subtracted the 10-Year Vegetative Season Average NDVI Map from each year's average NDVI map to produce a Change from the 10-Year Average NDVI Map for each year.

3.3.3. Assessment of Elm Tree Health

We divided the graphs into zones with the highest populations of elm trees—North Woods, 5th Avenue, Ramble, and Mall—to illustrate the health status of vegetation through NDVI mean values. These values reflect the vegetation health in each zone across various seasons. The metrics were analyzed on both an annual and seasonal basis.

We utilized in-situ data from the Central Park Conservancy, which provided point data on tree removal records. These records helped us understand the annual variations in tree removals. This, in turn, gave us insights into the full spread of elm bark beetles and their relationship to temperature and precipitation changes. We then generated a trend line based on the cumulative GDD and cumulative precipitation at the dates of the first tree removal each year. This step is essential for comprehending the relationship between GDD and precipitation at the onset of beetle spread. To assess the relationship between temperature (GDD) and precipitation, we categorized the study years into wet, normal, and dry years using three decades of precipitation data. This classification was based on precipitation quantiles including mean, standard deviation, dry threshold, and wet threshold. This approach enhances the understanding of how temperature and precipitation influence beetles spread around elm trees (Dreistadt & Lawson, 2014).

3.3.4. Elm Tree Health Assessment Using Remote Sensing

The validation of elm tree health was performed by combining the three datasets outlined in section 3.2.2 and running a logistic regression model. This model would calculate the relationship between pixel-level NDVI change and whether a pixel contained a known occurrence of Dutch Elm Disease. Furthermore, the model only ran the analysis on the pixel-level data until the first occurrence of Dutch Elm Disease for that year. The model was run to aggregate the data yearly.

For our project, we compared the Normalized Difference Vegetation Index (NDVI) values with observations from the Conservancy on tree health (e.g., Healthy vs. Unhealthy elm trees). We chose a systematic sampling method for our data (Wang et al., 2016). We employed 30m resolution NDVI pixel data for the same geographic area and time covered by the in-situ observations. Then, we organized our in-situ data with attributes such as latitude, longitude, tree health status, date of observations, and year. Each observation had a unique identifier, such as a tree ID or observation ID. Using ArcGIS Pro, we divided our study area into sections corresponding to the pixels of the Landsat 8 and 9 imagery. We assigned trees from the Conservancy's ground-referenced data to these VI pixels based on their geographical coordinates. Within each pixel, we sampled an equal number of trees to ensure a balanced representation across the study area (Lambert et al., 2013). After spatially joining the data, we created a contingency table to cross-tabulate NDVI classes and in-situ tree health status categories (Healthy, Unhealthy, removed trees). This table shows the distribution and association between NDVI classifications and observed tree health status. Finally, we performed statistical tests to assess the agreement between NDVI classes and in-situ tree health status. This process assesses the degree to which NDVI can effectively indicate tree health status based on comparisons with in-situ observations provided by the Conservancy.

4. Results & Discussion

4.1 Analysis of Results

4.1.1. *Land Surface Phenology*

Utilizing our NDVI time series, we created visualizations of the general phenological trends in Central Park (Figure 2). Central Park displays normal and predictable seasonality with NDVI values peaking in the months of June, July, and August and bottoming out in December, January, and February. This was a vital step in understanding the general foliar phenological trends for all of Central Park forest zones of concern and was later utilized for analyzing start and end of season dates. These dates would define the period within which we conducted tree canopy change analysis and DED spread factor analysis.

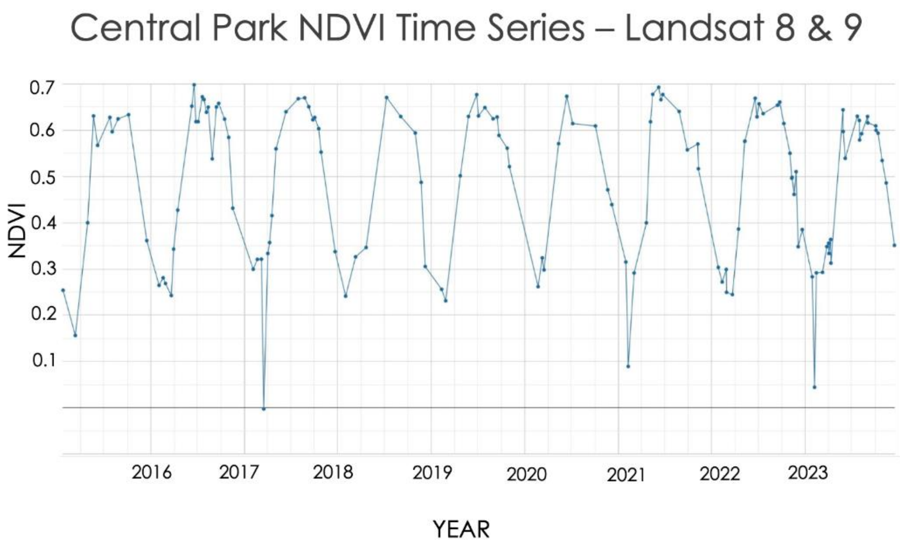

Figure 2. Landsat 8 & 9 Derived NDVI Time Series (2015-2023). The x-axis represents time in years, the y-axis displays NDVI values. The NDVI values shown are based on data for Central Park's entire tree canopy cover.

After applying the Generalized Additive Model (GAM) analysis to our NDVI data, we created bar charts to visualize the start and end of the growing season, as well as the peak vegetation period for each year (Figure 3). Our analysis revealed that the start of the growing season was relatively consistent every year, typically beginning between mid-April and mid-May. This variability in the start date could be influenced by differences in spring temperatures and precipitation patterns, which affect the timing of vegetation growth. Similarly, the end of the vegetative season (vegetation greenness) consistently occurred in September. However, the exact timing within this month may also be influenced by local climatic conditions. Peak vegetation, which represents the period of maximum NDVI values, shifted yearly. This shift is likely attributable to variations in early summer weather conditions, such as temperature and precipitation, which can affect the rate of plant growth and the timing of peak biomass accumulation.

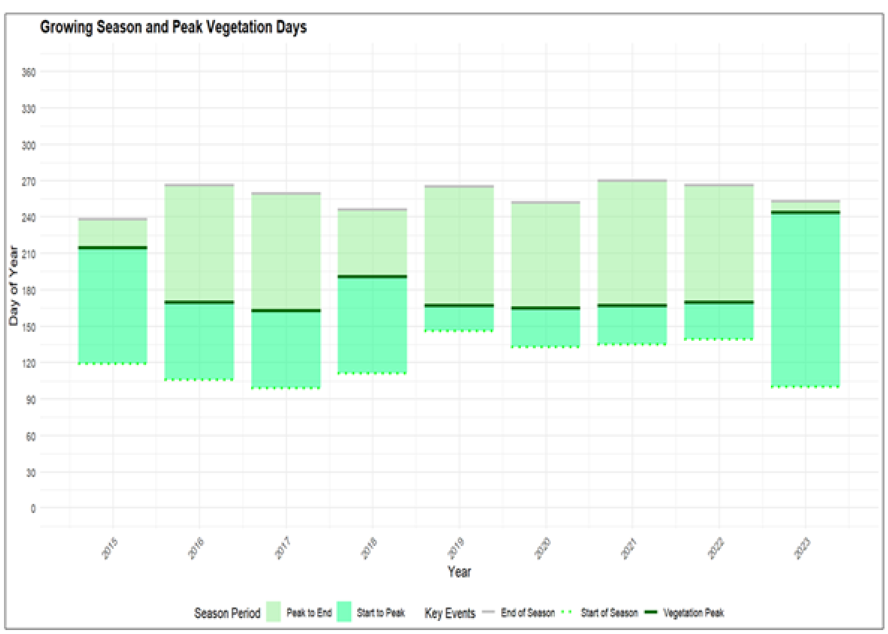

Figure 3. Key events for the growing season in Central Park from 2015 to 2023. The x-axis displays time in years and the y-axis displays time in days. The colors in the graph represent different key times within the

season for each year and when they occurred (Mint = start of season to peak of season, Olive = peak of season to the end of the season, Black Line = peak of the season, Grey Line = end of season).

These observations suggest that annual fluctuations in weather patterns play a significant role in determining the growing season's characteristics (Ye et al., 2021). The GAM analysis and visualizations provide valuable insights into the temporal variability of the growing season and peak vegetation, highlighting the influence of weather factors on trees' canopy greenness and health status (Seasonal Vegetation Dynamics).

4.1.2 Tree Canopy Change

We identified hotspots in the park using each year's Change from the 10-Year Average NDVI Map. In the 2023 vegetative season, for example, a cluster of low-NDVI cells in the northeast section of the park highlighted the Harlem Meer Center construction project's impact on tree canopy health (Figure 4). The maps also indicated which parts of the park flourished in certain years. In 2023, for example, there was an increase in NDVI in the southern section of the park. When comparing the maps to the Conservancy's record of DED occurrences, however, the DED occurrences did not necessarily correspond to a decrease in NDVI on the maps in the occurrence locations. This was not unexpected, since the maps combined multiple acquisitions from the vegetative season. This method was better suited for observing general phenomena rather than precise DED events.

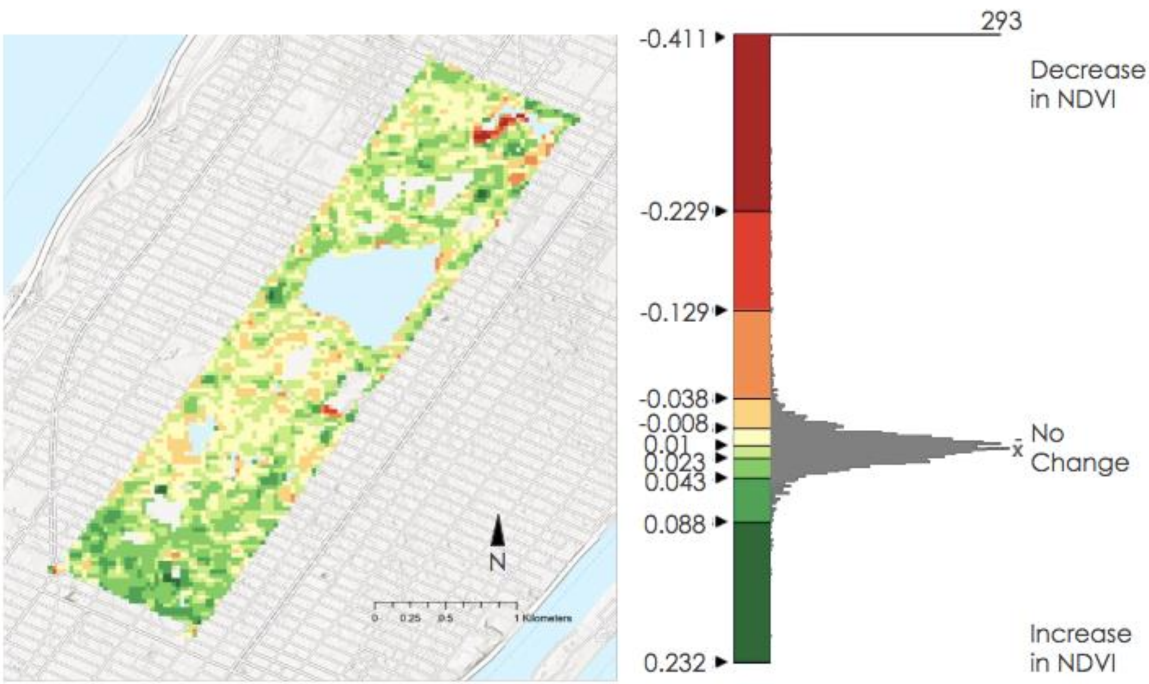

Figure 4. 2013 Vegetative Season NDVI Compared to the 10-Year Average. Dark red values represent a decrease in NDVI, and dark green represents an increase in NDVI. The grey histogram displays the number of pixels at a certain recorded NDVI value.

4.1.3. Assessment of Elm Tree Health

The time series data was then restructured to create monthly zonal averages (Figure 5). This analysis was performed across the entire time to see the rate of seasonal onset. Finally, the NDVI data was used with precipitation data to gain a deeper understanding of the start and end of season in the park. On average, all zones of the park experience an initial season onset between March and April with a steeper positive NDVI slope from April to May. Finally, between June and October, the NDVI value plateaus before dropping down in the winter months.

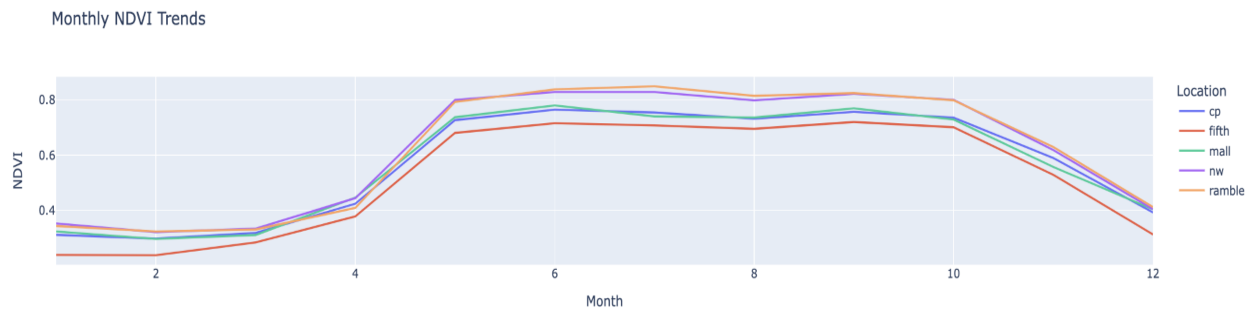

Figure 5. Monthly NDVI Trends Derived from Mean NDVI Landsat 8 & 9 Data (2014-2023). The x-axis represents months of the year, the y-axis represents mean NDVI values across the 2014-2023 study period to display monthly average NDVI. Each color represents a different zone within the park (Blue = All Central Park, Red = Fifth Avenue, Green = The Mall, Purple = The Northwoods, Orange = The Ramble).

We created the final visualization by utilizing the Landsat-derived NDVI time series (Figure 6). We analyzed the NDVI values in specific zones of the park and the entire Central Park canopy. This provided the team with greater insight into how the zones differ from each other and the overall park. The data shows that Ramble and North Woods have typically higher NDVI values than the rest of the park. Furthermore, the Fifth Avenue NDVI values are consistently lower than the other zones. This was expected as Fifth Avenue has a comparatively sparse canopy. In contrast, the Ramble and the North Woods are considered naturally forested areas of the park and have dense forest canopies. The data displayed in Figure 5 affirmed our hypothesis that NDVI would differ between zones in the park.

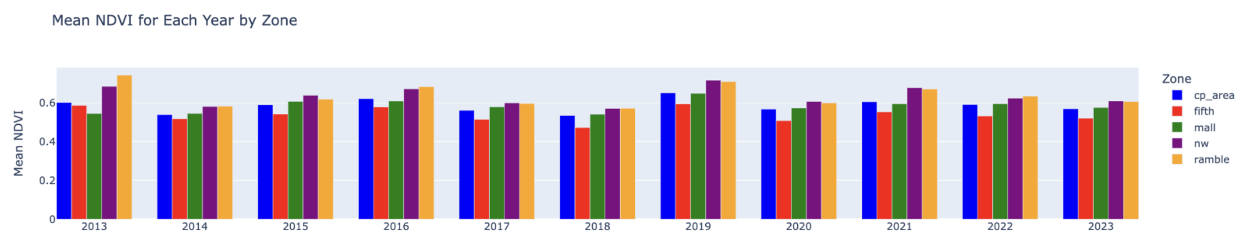

Figure 6. Mean Yearly NDVI Comparison Between Park Zones. The x-axis represents time in years, the y-axis represents mean NDVI values for the vegetative season of that year. Each color represents a different zone within the park (Blue = All Central Park, Red = Fifth Avenue, Green = The Mall, Purple = The Northwoods, Orange = The Ramble).

We collected data on cumulative growing degree days and cumulative precipitation from January 1st until the first tree removal date each year to analyze the relationship between precipitation and temperature. The trend line we created showed a negative correlation, indicating that higher cumulative GDD was associated with lower cumulative precipitation during the peak spread of beetles each year (Figure 7).

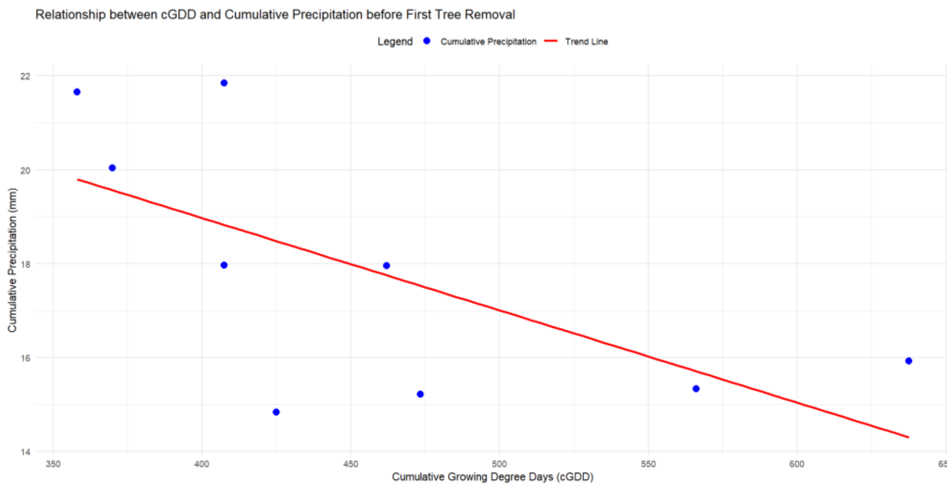

Figure 7. The relationship between cumulative Growing Degree Days and cumulative precipitation before the first tree removal of the season.

To further understand this relationship, we categorized the study years as wet, normal, or dry based on the threshold method (Table 2). We then correlated these classifications with the beetle spread dates. The results showed that beetles spread at higher GDD levels during dry years, while they spread at lower GDD levels during wet years (Figure 8).

Table 2

Years classification (wet, normal, or dry) based on precipitation data

Measure	Value
Precipitation - Mean	0.142
Precipitation - Standard Deviation	0.024
Dry Threshold	0.117
Wet Threshold	0.166

These findings demonstrate that elm bark beetles require not just temperature but also sufficient moisture to thrive and spread effectively. Therefore, both temperature and precipitation are crucial factors to consider when studying or predicting pest spread. This insight highlights the importance of incorporating both variables in models to accurately predict beetle infestations and their potential damage to elm trees.

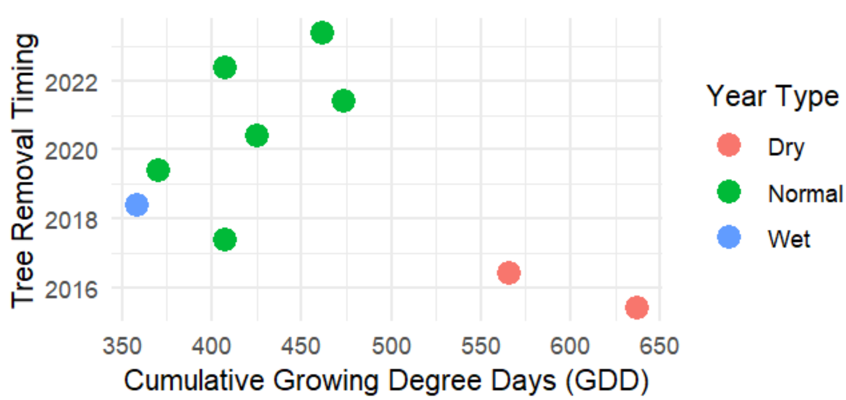

Figure 8. Cumulative Growing Degree Days versus the date of the first tree removal in the season. The x-axis displays cumulative growing degree days, the y-axis shows time in years. Each point's color indicates the precipitation of that year (Red = Dry, Green = Normal, Blue = Wet).

4.1.*4. Elm Tree Health Validation Using Remote Sensing*

We ran a logistic regression model to better understand the correlation and apparent accuracy of using Landsat 30-meter NDVI data for detecting DED occurrences. The first step in this process was to derive pixel-level NDVI data for the entire study period. As shown in Figure 9, some outlying pixels experienced drops in NDVI values between images that were available from Landsat 8 and 9. We utilized this data to detect small NDVI fluctuations that might be logically consistent with an occurrence of Dutch Elm Disease.

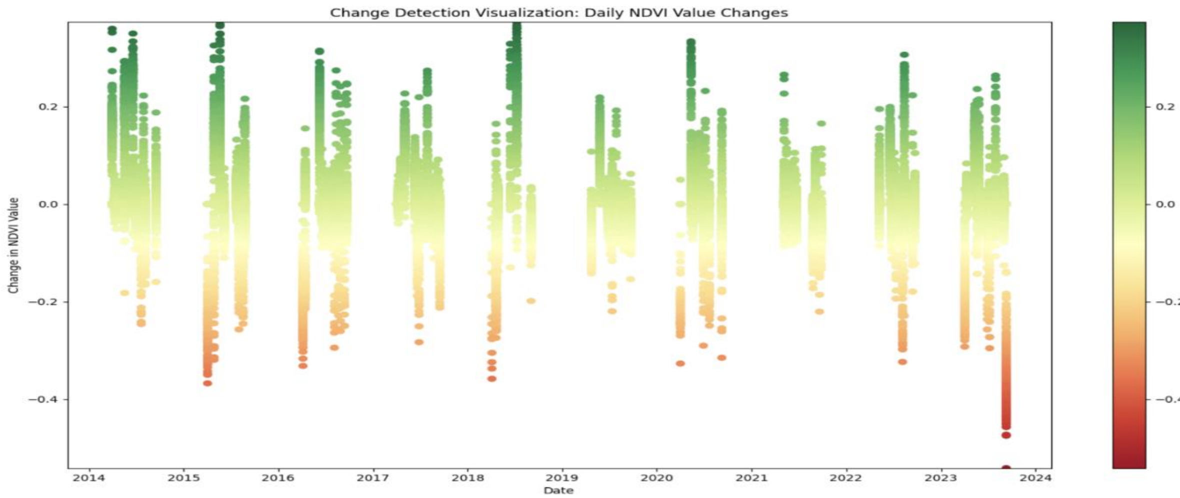

Figure 9. Pixel Level NDVI Change Between Satellite Acquisitions. The x-axis displays time in years, the y-axis shows the change in NDVI value for an individual pixel between satellite imagery acquisitions. Each point represents one pixels NDVI value at the recorded date.

Utilizing the combined dataset containing tree locations, tree activities, and the pixel-level NDVI data, we ran a logistic regression model which validated the ability to detect whether a tree did or did not have DED. The logistic regression model utilized three columns within the final dataset: NDVI value, the date on which a DED occurrence happened, and a True/False value indicating whether the pixel did or did not have DED infected. The logistic regression model runs a statistical modeling analysis to determine the relationship between these variables. These findings were vital to understanding whether our methodology was effective or not. Based on the precision metric, the model showed a high degree of success (71%) in detecting unhealthy tree canopies (Table 3). In contrast, the model yielded a lower precision rate (41%) in predicting healthy tree canopies. To derive the logistic regression model, 307 elm tree pixels were analyzed, 218 pixels were infected, and 89 were healthy trees.

Table 3

Logistic Regression Model Results

Canopy Type	Precision	Recall	F-1 Score	Observations
Healthy Canopies	0.37	0.47	0.41	89
Unhealthy Canopies	0.76	0.67	0.71	218

4.2 Errors & Uncertainties

The errors and uncertainties of this project can be split into two categories: those inherent to the phenomena being studied, and those inherent to the data acquired and analysis performed. In the first category, DED is spread not only by insect vectors like the elm bark beetle. According to the Conservancy's tree care team, the disease is more commonly spread in Central Park through root grafts in the soil between infected and healthy trees, which cannot be detected directly using satellite data. This study also cannot account for the natural variability in forest foliage phenology (e.g., for deciduous forest tree species). NDVI was used as a proxy for indicators of forest phenology and tree canopy health, yet minor changes in this index due to natural variability or variability in satellite measurements may be on the same order of magnitude as changes in this index (due to certain effects of DED, such as damage to foliage on certain branches of a given tree). Furthermore, Landsat 7, 8, and 9's spatial resolution of 30m can create inconsistencies in the data as one pixel can be a mixture of multiple tree crowns, containing both healthy and unhealthy trees. In addition, times of low temporal resolution can negatively impact the prediction of start and end of season. Finally, it is critical to recognize that this study attempted to detect a disease in an area in which prevention and treatment measures for that disease are actively being performed. Thanks to the Conservancy's stewardship, the elm trees of Central Park are cared for unlike any other collection of elms in the world, and this study does not account for the effects of this maintenance on the detection of DED. This might be addressed by comparing Central Park to a similar area containing elms that were not maintained or treated and had known infections of DED.

Secondly, there are considerations to understand about the data acquired and the analysis performed. Our validation showed that in some cases Landsat 8 and 9 can detect DED impacts with a moderate degree of accuracy for pixels that contain sufficient elm crown coverage to dominate the pixel. Yet, the uncertainties due to natural variability could possibly be reduced using satellite imagery that has higher spatial and/or temporal resolutions. The validation analysis was performed after joining datasets based on geographical coordinates. Slight differences in the coordinates of the tree sites and tree tasks datasets were rounded to the fourth decimal so that they could be joined, and then this was joined by coordinates to the NDVI pixel data. These differences and the rounding could have introduced errors, in which tree sites and tasks were incorrectly joined to a pixel adjacent to the correct one. Finally, there are anomalies to note in 2017, 2021, and 2023 (Figure 2). These anomalies could be due to data interference, cloud cover, or human activity leading to unexpectedly low NDVI averages which decrease the accuracy of the validation. Finally, the total dataset of tree points had a higher number of unhealthy trees as the disease is pervasive throughout the park impacting trees at various points in their life cycles. Furthermore, detecting subtle NDVI changes of healthy trees as management strategies and the disease are simultaneously spreading may create inconsistencies in assigning a tree a healthy or unhealthy value. This could be one reason as to why the accuracy for detection is lower in healthy tree canopies.

4.3 Feasibility & Partner Implementation

We found through our work that NASA Earth observations can be used to quantify NDVI change and certain phenological metrics for the Central Park overstory tree canopies. Secondly, through our validation, we determined that Landsat 8 and 9 can identify unhealthy tree-dominated pixels that have DED infections, and ones that do not, though the accuracy in doing so is far from perfect. The validation results provided some evidence that it is technically possible to use remote sensing methods to generate a predictive model for DED detection, which is of interest to the partner. However, we could not fully explore this potential due to time limitations. We hope that these insights into phenological timing and change, DED spread factors, and DED detection using remote sensing can all aid in the Conservancy's pursuit of proactive tree canopy evaluations and DED management planning.

5. Conclusions

Based on our study, we determined that NASA Earth observations enable feasible methods for assessing urban forest phenology and detecting occurrences of Dutch Elm Disease through pixel-level NDVI analysis. While the potential for errors in our methodology and uncertainties within the data exist, our team's work outlines a clear and employable method for aiding our partner's concerns about urban tree health monitoring and management. In the future, this study can be further developed through consideration of several other methods that include higher temporal or spatial resolution satellite and/or aerial data, more precise recording of DED occurrences on behalf of the conservancy, and a unified method of geospatially referencing trees as point data that is consistent between datasets. These methods could possibly improve upon the results of our feasibility study considerably and are valid avenues for further exploration.

6. Acknowledgements

We express our gratitude to our partners Sean Cameron, a DEVELOP alumnus, and Yanina Kupava for providing our team with invaluable knowledge about the park and the Conservancy's management practices. We are thankful to have been guided by the expertise of our science advisors Dr. Mehdi Heris, Joseph Spruce, and Dr. Kenton Ross, as well as our science manager Amanda Clayton. Our team is grateful for all the support provided by our lead Ella Haugen throughout the term. We would also like to thank Maya Hall, Fellow at DEVELOP Ames, for reviewing this paper.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Aeronautics and Space Administration.

This material is based upon work supported by NASA through contract 80LARC23FA024.

7. Glossary

Diameter at Breast Height (DBH) – A method of measuring a tree trunk's diameter, measuring at the observer's breast height.

Digital Elevation Model (DEM) – A band in LiDAR used to model topography.

Dutch Elm Disease (DED) – A devastating fungal disease that affects elm trees, caused primarily by the pathogens *Ophiostoma ulmi* and *Ophiostoma novo-ulmi*.

Earth observations – Satellites and sensors that collect information about the Earth's physical, chemical, and biological systems over space and time.

Generalized Additive Model (GAM) analysis – A flexible statistical method used to model complex relationships between a response variable and one or more predictor variables.

Global Historical Climatology Network-Daily (GHCN) – a database of daily climate summaries from land surface stations around the globe.

Growing Degree Days (GDD) – Cumulative heat units received during the vegetative or growing season.

Integrated pest management (IPM) – A practical and eco-friendly pest management strategy that uses a blend of sensible practices.

Light Detection and Ranging (LiDAR) – An active form of remote sensing that measures distances between the sensor and the surface of the Earth using light from a laser.

Normalized Difference Vegetation Index (NDVI) – A remote sensing metric used to assess and monitor vegetation health and density.

Quality Assurance (QA) – A band used with the Landsat collection to mask clouds.

8. References

Banskota, A., Kayastha, N., Falkowski, M. J., Wulder, M. A., Froese, R. E., & White, J. C. (2014). Forest Monitoring Using Landsat Time Series Data: A Review. *Canadian Journal of Remote Sensing*, *40*(5), 362–384. https://doi.org/10.1080/07038992.2014.987376

Bounds M., Feller M., Greenfeld J., Heaviland M., Pregitzer C., Wenskus T. (2014). *Guidelines to urban forest restoration*. New York City Department of Parks and Recreation. New York, NY. https://www.nycgovparks.org/pagefiles/84/guidelines-to-urban-forest-restoration.pdf

Broderick, T. (2024, April 9). *The Fight to Save Central Park's Iconic American Elm Trees*. Central Park Conservancy. https://www.centralparknyc.org/articles/american-elms-central-park-mall

Caparros-Santiago, J. A., Rodriguez-Galiano, V., & Dash, J. (2021). Land surface phenology as indicator of global terrestrial ecosystem dynamics: A systematic review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *171*, 330–347. https://doi.org/10.1016/j.isprsjprs.2020.11.019

Central Park Conservancy. (2024). *About Us*. Central Park Conservancy. Accessed on 06/25/2024 https://www.centralparknyc.org/about

Central Park Conservancy. (2024). *Park History*. Central Park Conservancy- Park History. Accessed on 06/24/2024 https://www.centralparknyc.org/park-history

Copeland, C., Harper, R., Brazee, N., & Bowlick, F. (2022). A review of Dutch elm disease and new prospects for Ulmus americana in the urban environment. *Arboricultural Journal*, *45*, 1–27. https://doi.org/10.1080/03071375.2022.2082177

Dean, A., & Hodgson, E. (2022). *Growing Degree Days for Insect Pests | Integrated Crop Management—University of Iowa*. https://crops.extension.iastate.edu/encyclopedia/growing-degree-days-insect-pests

Dreistadt, S. H., & Lawson, A. B. (2014). *Elm Leaf Beetle Management Guidelines—UC IPM*. https://ipm.ucanr.edu/PMG/PESTNOTES/pn7403.html

Guo, D., Wang, C., Zang, S., Hua, J., Lv, Z., & Lin, Y. (2021). Gap-Filling of 8-Day Terra MODIS Daytime Land Surface Temperature in High-Latitude Cold Region with Generalized Additive Models (GAM). *Remote Sensing*, *13*(18), Article 18. https://doi.org/10.3390/rs13183667

Hargrove, W. W., Spruce, J. P., Gasser, G. E., & Hoffman, F. M. (2009). Toward a national early warning system for forest disturbances using remotely sensed canopy phenology. *Photogrammetric Engineering & Remote Sensing*, *75*, 1150–1156.

Lambert, J., Drenou, C., Denux, J.-P., Balent, G., & Cheret, V. (2013). Monitoring forest decline through remote sensing time series analysis. *GIScience & Remote Sensing*, *50*(4), 437–457. https://doi.org/10.1080/15481603.2013.820070

Landsat 8. (2021, November 30). Landsat Science. https://landsat.gsfc.nasa.gov/satellites/landsat-8/

Landsat 9. (2021, November 30). Landsat Science. https://landsat.gsfc.nasa.gov/satellites/landsat-9/

Li, X., Chen, W. Y., Sanesi, G., & Lafortezza, R. (2019). Remote Sensing in Urban Forestry: Recent Applications and Future Directions. *Remote Sensing*, *11*(10), Article 10. https://doi.org/10.3390/rs11101144

Murray, M. (2020). *Using Degree Days to Time Treatments for Insect Pests- University of Utah.* https://extension.usu.edu/planthealth/research/degree-days

Neyns, R., & Canters, F. (2022). Mapping of Urban Vegetation with High-Resolution Remote Sensing: A Review. Remote Sensing, 14(4), Article 4. https://doi.org/10.3390/rs14041031

Office of Technology and Innovation (OTI). (2022). *Land Cover Raster Data (2017) – 6in Resolution | NYC Open Data.* Accessed on 06/25/2024 https://data.cityofnewyork.us/Environment/Land-Cover-Raster-Data-2017-6in-Resolution/he6d-2qns/about_data

Staley, D. (2022). Modern urban forestry for modern cities: Technology challenges and opportunities for the remote sensing of urban forests. Arboriculture & Urban Forestry, 48(2), 147-163. https://doi.org/10.48044/jauf.2022.012.

Sutliff, A. (2023, March 13). Landscapes of Central Park. Central Park Conservancy. https://www.centralparknyc.org/articles/landscapes-of-central-park

Wang, S., Li, X., Ge, Y., Jin, R., Ma, M., Liu, Q., Wen, J., & Liu, S. (2016). Validation of Regional-Scale Remote Sensing Products in China: From Site to Network. *Remote Sensing*, *8*(12), Article 12. https://doi.org/10.3390/rs8120980

Ye, W., van Dijk, A. I. J. M., Huete, A., & Yebra, M. (2021). Global trends in vegetation seasonality in the GIMMS NDVI3g and their robustness. *International Journal of Applied Earth Observation and Geoinformation*, *94*, 102238. https://doi.org/10.1016/j.jag.2020.102238